September, 2012

Reading the Literals:
Searching For Expressions in Text on Death Certificates In Montana, 2003-2010

Cody Custis, MS, Epidemiologist, MHDDS
Bruce Schwartz, MA, MPA, Lead Vital Statistics Epidemiologist

When an individual dies in Montana, a medical certifier fills out cause-of-death information and submits it to Montana Office of Vital Statistics (OVS) Data Acquisition staff. Since the 1989 revision of the U.S. Standard Certificate of Death, this information has consisted of four lines for the causes of death; a second part for other conditions which contributed to, but did not directly result in the underlying cause of the death; and, for a death attributed to injury, a narrative description of the injury.[1] These fields, called "the literals" in this summary, are transmitted to the National Center for Health Statistics (NCHS); NCHS returns multiple ICD codes based on the literals,[2] which staff of the Vital Statistics Analysis Unit append to the legal and demographic portions of the record. Ninety percent of records are coded by a computer program developed by NCHS; the remaining 10% cannot be automatically coded and are manually coded by trained nosologists. Montana sends death certificate records to the NCHS on a bi-weekly basis; NCHS returns cause of death codes with a delay of three weeks, although the delay may be longer for rare causes of death, which are more likely to require manual review. In 2003, Montana implemented a new revision of its death certificate.[3] Prior to 2003, Montana's vital statistics database made ICD codes, but not literals, available to vital statistics and other analysts in the Montana Department of Public Health and Human Services (DPHHS). Because OVS began to collect death records electronically in the same year, it was possible to make the complete death certificate as filed, including the literals, available in the vital statistics database.

This report highlights three examples of the value of examining these literals more closely in the public health analysis of death certificates. The first used literals to investigate a suspected outbreak of infectious disease quickly, the second used literals to assess the contribution of suspected risk factors to infant deaths, and the third used literals to identify a specific demographic subset of decedents who committed suicide. Information about deaths occurring to Montana residents out of state is shared through interstate agreements, but records from out of state do not consistently include the literals. Therefore, this report focuses on deaths that occurred in Montana, regardless of the state of residence of the decedents.

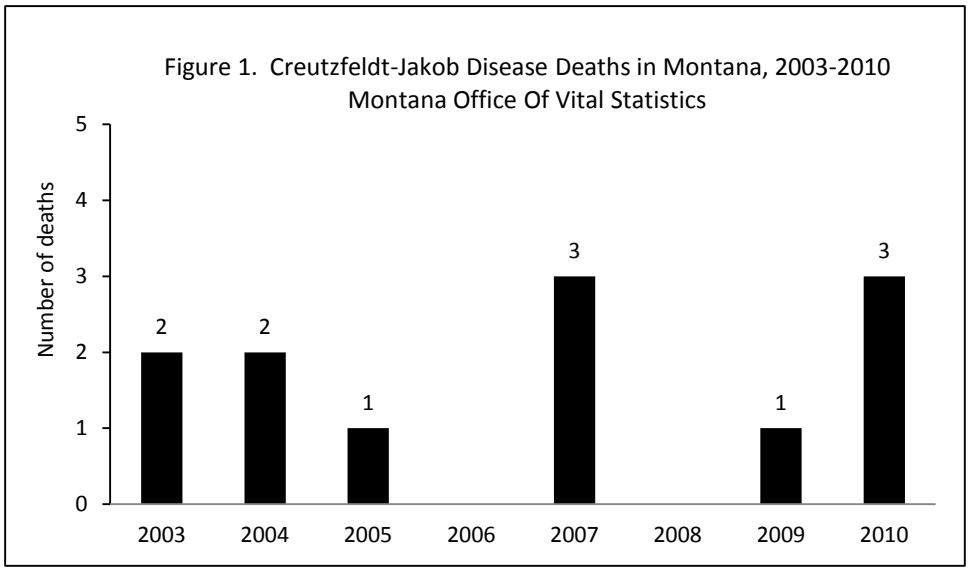
Creutzfeldt-Jakob Disease

Classic sporadic Creutzfeldt-Jakob Disease (sCJD) is a fatal degenerative disease of the brain which affects approximately one in a million people a year worldwide; 85% of sCJD cases are sporadic and the remainder appears to be familial.[4,5] Variant CJD (vCJD) is probably related to Bovine Spongiform Encephalopathy

[1] http://www.cdc.gov/nchs/data/series/sr_04/sr04_028.pdf
[2] Moriyama IM, Loy RM, Robb-Smith AHT. History of the statistical classification of diseases and causes of death. Rosenberg HM, Hoyert DL, eds. Hyattsville, MD: National Center for Health Statistics. 2011. Available at http://www.cdc.gov/nchs/icd.htm
[3] http://www.cdc.gov/nchs/nvss/vital_certificate_revisions.htm
[4] http://www.ninds.nih.gov/disorders/cjd/detail_cjd.htm

(BSE), which humans may contract by eating infected beef neural tissue. In 2003, a possible cluster of fatal human degenerative neurologic disease was reported in Wisconsin among men who regularly consumed wild game together, raising concern that they may have contracted a transmissible spongiform encephalopathy from deer or elk suffering from chronic wasting disease (CWD), endemic to wild or captive herds in several Midwest and Mountain states. These conditions are all attributed to infection with prions, atypical proteins that preferentially attack brain and other neural tissue, and that cause the body's own proteins to take on the atypical form.[6] Symptoms are progressive, in some cases rapidly so, and infection is always fatal.[6] The ICD-10 mortality code A81.0 does not distinguish between sCJD, vCJD, BSE, CWD, and other presumptive prion diseases.

Montana typically has one or two deaths per year from CJD (Figure 1). In 2011, a suspected cluster of three CJD cases, from a single county and diagnosed within a relatively short period of time, was reported to the Montana DPHHS.[7] As part of the cluster investigation, we searched recent death certificates for other potential cases. We developed algorithms to find additional CJD cases by referring to literals used on CJD death certificates from 2003 to 2010. Spelling variations of 'CREUTZFELDT' and 'JAKOB' were particularly challenging. No additional cases of CJD were discovered and the suspected cluster was attributed to chance.



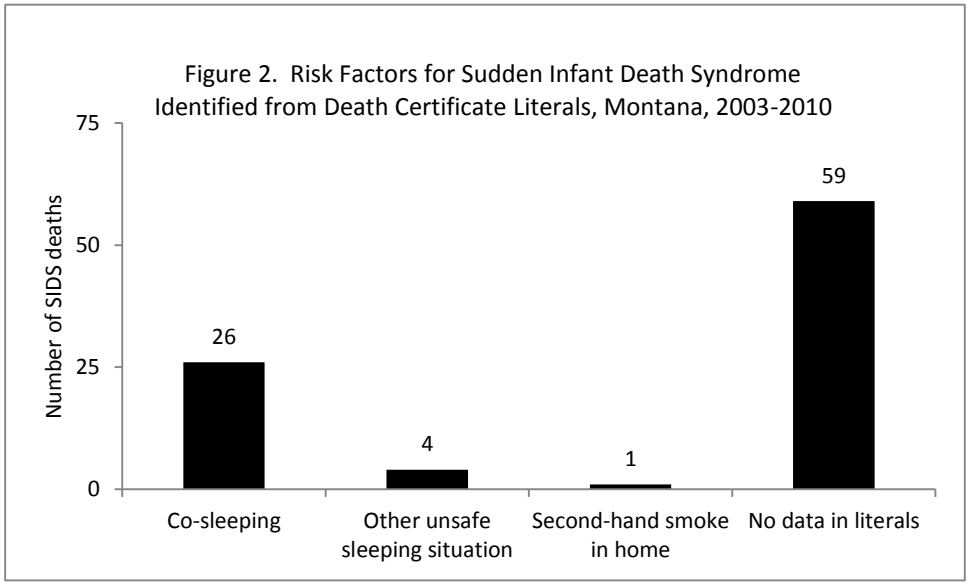
This technique is also applied to death certificates during regular monthly processing by the Montana OVS Data Acquisition staff. Because of the relatively long turn-around time for cause-of-death codes from NCHS, our staff uses this technique to locate records of deaths from several communicable diseases more quickly, and reports results to the Communicable Disease Epidemiology Program of DPHHS. Epidemiologists need to be able to respond to point source outbreaks quickly in order to control disease and to and respond to public concern. The delay in receiving coded record back from NCHS is sometimes too long when preventive action needs to be taken promptly. By searching the literals, Vital Statistics is able to provide information on key communicable disease death in a timely manner.

[5] http://www.cdc.gov/ncidod/dvrd/cjd/
[6] http://www.cdc.gov/ncidod/dvrd/prions/
[7] http://www.dphhs.mt.gov/publichealth/preventionopportunities/2012/2012-06MPH.pdf

SIDS Deaths

Sudden Infant Death Syndrome (SIDS) is an unexpected death of an infant for which an autopsy and complete investigation of the circumstances of death do not yield a clear explanation.[8] There are theories about the physiological basis of SIDS deaths, but these remain largely speculative.[9] However, several well-established risk factors that have been the basis of successful prevention campaigns.[10]

We examined Montana infant deaths with SIDS (ICD-10: R95) as an underlying cause of death to look for information on risk factors in the literals. There were 90 SIDS deaths in Montana between 2003 and 2010. We searched the literals for terms indicating unsafe infant sleeping practices such as putting the infant to sleep on an inappropriate bed or surface; surrounding the infant with soft or loose bedding; bed-sharing or co-sleeping with the infant; and not putting the infant to sleep on his or her back. The search yielded 26 deaths with comments about co-sleeping with the infants, and an additional five deaths with comments about other unsafe circumstances (not on back, inappropriate bed or bedding, second-hand smoke in home) (Figure 2). The remaining 59 records contained no literals so our ability to estimate the prevalence of risk factors from the infant death certificates proved to be very limited. SIDS and all other fetal, infant, and childhood deaths are reviewed by Fetal, Infant, and Child Mortality Review Committees in Montana,[11] which are in a position to provide much more thorough assessment of risk factors.



Suicide Among Prison Inmates

According to a report to the Montana legislature, Montana has an inmate suicide rate five times the national average.[12] In response to an inquiry about inmate suicide, we developed algorithms to search for

[8] http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0002533/
[9] Kadhim H et al. 2010. *Neurosci Lett* 480:122-126; Paterson DS et al. 2006 *JAMA* 296:2124-2132.
[10] American Academy of Pediatrics Policy Statement. 2011. *Pediatrics* 128:1030-1039.
[11] http://data.opi.mt.gov/bills/mca_50_19_4.htm
[12] Law and Justice Interim Committee Report, June 18, 2010 (http://www.krtv.com/files/jailsuicide.pdf)

suicide deaths among prison inmates.[13] We limited the analysis to records with suicide as the underlying cause of death (ICD-10: X60-X84). We searched cause of death literals, descriptions of injuries, occupation and industry fields, and address fields for *'prison', 'jail', 'convict',* and *'incarcerate.'* We found seven inmate suicides in Montana from 2003-2010 using the initial terms. After reviewing those records, we added other terms seen on inmate death certificates: *'inmate'*, *'cell'*, and *'detention'* and an additional field, place of injury. This increased the list to 32 possible inmate suicides; five were later excluded after review because the technique can lead to the flagging of false positive cases. For example, one decedent was '*incarcerate*d in *prison* most of his life' but committed suicide at his residence; one had 'small *cell* lung cancer;' one 'wrapped *cell* phone charger around neck;' one was a '*cell* tower installer;' and one was a '*detention* officer.' Table 1 shows that certain key words were found in specific fields. However, it is advisable to search all likely fields for each term because certifiers may write useful information in any available space.

Table 1. Selected Literal Expressions Indicating Inmate Status

Literal(s)	Contained in Field(s)	Number Of Records
COUNTY *JAIL*	Address	1
JAIL CELL	Description Of Injury	1
INMATE, *PRISON*, *PRISON*	Occupation, Industry, Address	1
INMATE, *PRISON*	Occupation, Address	1
CELL DOOR	Description Of Injury	1
DETENTION CENTER, *DETENTION* CENTER	Description Of Injury, Address	1
INMATE	Occupation, Industry	1
JAIL PANTS TO THE *JAIL CELL*	Description Of Injury	1
INMATE AWAITING TRIAL IN A *DETENTION* FACILITY	Cause Of Death, Description of Injury	1
DETENTION CENTER	Address	1
CELL BED	Description Of Injury	1
HOLDING *CELL*	Description Of Injury	1
DETENTION CENTER *CELL*	Description Of Injury	1
INMATE	Occupation, Industry	1
INMATE WHICH PLEAD GUILTY, *DETENTION* CENTER	Description Of Injury, Address	**1**
CELL SINK, *DETENTION* CENTER	Description Of Injury, Address	1
BUNK BED IN HIS *CELL*	Description Of Injury	1
INMATE, *DETENTION* CENTER	Description Of Injury	1
DETENTION CENTER	Place Of Injury	4
PRISON	Place Of Injury	1
JAIL	Place Of Injury	1
DETENTION FACILITY	Place Of Injury	3

The literals let us quickly identify classes of individuals of interest. We found records for 27 inmate suicides with a simple programming routine. We used the first cases we identified to expand the search with additional terms and fields associated with inmate status. For more complicated searches, there might be several iterations of this process. However, the technique is never likely to be exhaustive. We might have found

[13] Montana OVS suicide numbers may vary from those reported by Montana correctional facilities, which are required to directly submit inmate deaths to the Bureau of Justice Statistics

more inmate suicides if we had searched for spelling variations and typographical errors which are common in the literals. It is nearly impossible to anticipate all such errors. This example also underscores the need for manual review of records to verify that the individuals actually were inmates, because '*cell*' might refer both to detention cells and also to cellular communications or cells in the body. Likewise, we had to manually review records to identify individuals who worked in the detention system as guards and officers, but were not incarcerated.

Searching literals will not find suicides at Montana detention facilities that do not have specified incarceration-related expressions in the fields searched. We know, based on information from other sources, that our enumeration of inmate suicides from reviewing literals might have missed as many as one third of the cases. Our ability to identify inmate suicides depends on certifiers of death including text that we can associate with inmate status.

Conclusion

Searching literal cause of death and description of injury fields has expanded the usefulness of death records beyond ICD codes. The codes are useful for classifying underlying cause of death and have been the basis for cause-of-death tabulations for several decades, but are inadequate as the sole source for timely identification and investigation of potential clusters of deaths from infectious disease. Likewise, ICD codes do not identify subgroups of deaths of interest, such as prison inmates. Literals are also useful for examining risk factors and prevention opportunities, such as the association between sleeping circumstances in SIDS deaths.

There is, however, a major limitation of literals: they depend on certifiers completing as many of the lines in the cause-of-death section of the death certificate, and including descriptive information in the sections for contributing causes of death and circumstances of injuries on death certificates. In the example of prisoner suicides, ancillary information leads us to believe there are many more suicides in custody than we were able to ascertain from death certificates because certifiers did not record key words that would allow us to recognize prisoner status. In the example of SIDS deaths, two thirds of the cases coded as SIDS did not have any pertinent information at all in the literal portions of the certificates.

Absence of information in the literals does not mean "No." *It is missing data and must be treated as such.* Although Vital Statistics analysts now have access to the full content of death certificates in electronic format because certificates are filed on-line, the great majority of death certificates may lack relevant literals.

For information about Montana Vital Statistics, please contact Bruce Schwartz, Lead Vital Statistics Epidemiologist, Office of Epidemiology and Scientific Support, (406) 444-1756 or bschwartz@mt.gov
This document was published in electronic form only. Alternative formats of this document will be provided on request.
Please visit our website at http://www.dphhs.mt.gov/publichealth/epidemiology/index.shtml